Multiple Imputation with Angular Covariates

Imputing Incomplete Angular Data with Projected Normal Regression

Benjamin Stockton

University of Connecticut, Storrs, CT

3/21/2023

Table of Contents

Overview of the Presentation

- The projected normal distribution and multiple imputation

Overview of the Presentation

- The projected normal distribution and multiple imputation
- Complications with imputing angular data

Overview of the Presentation

- The projected normal distribution and multiple imputation
- Complications with imputing angular data
- A novel projected normal-based imputation method for angular data

Overview of the Presentation

- The projected normal distribution and multiple imputation
- Complications with imputing angular data
- A novel projected normal-based imputation method for angular data
- Simulation results

Overview of the Presentation

- The projected normal distribution and multiple imputation
- Complications with imputing angular data
- A novel projected normal-based imputation method for angular data
- Simulation results
- An illustrative analysis of US adult asthma rates' connection to pollution and wind

Table of Contents

Review of Multiple Imputation

- Let $\mathbf{X} = (X_1, \ldots, X_p, Y)$ be complete and $\Theta = (\theta, \cos\theta, \sin\theta)$ be partially observed, $\epsilon \sim N_n(0, \sigma^2 I_n)$, and

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \beta_{p+1}\cos\theta + \beta_{p+2}\sin\theta + \epsilon$$

Multiple Imputation Procedure [Rubin, 1987] :

1. **Impute** - Use an imputation method g to impute Θ_{mis} by $\dot{\theta} = g(\mathbf{X}, \Theta_{obs})$ or $(\cos\dot{\theta}, \sin\dot{\theta})' = g(\mathbf{X}, \Theta_{obs})$ M times to create M completed data sets

Review of Multiple Imputation

- Let $\mathbf{X} = (X_1, \ldots, X_p, Y)$ be complete and $\Theta = (\theta, \cos\theta, \sin\theta)$ be partially observed, $\epsilon \sim N_n(0, \sigma^2 I_n)$, and

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \beta_{p+1}\cos\theta + \beta_{p+2}\sin\theta + \epsilon$$

Multiple Imputation Procedure [Rubin, 1987] :

1. **Impute** - Use an imputation method g to impute Θ_{mis} by $\dot{\theta} = g(\mathbf{X}, \Theta_{obs})$ or $(\cos\dot{\theta}, \sin\dot{\theta})' = g(\mathbf{X}, \Theta_{obs})$ M times to create M completed data sets
2. **Analyze** - For each completed data $(\mathbf{X}, \Theta_{obs}, \Theta_{mis}^{(m)})$ for $m = 1, \ldots, M$, estimate β using least squares to collect $\hat{\beta}^{(m)}$ and $U^{(m)} = se(\hat{\beta}^{(m)})^2$

Review of Multiple Imputation

- Let $\mathbf{X} = (X_1, \ldots, X_p, Y)$ be complete and $\Theta = (\theta, \cos\theta, \sin\theta)$ be partially observed, $\epsilon \sim N_n(0, \sigma^2 I_n)$, and

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \beta_{p+1}\cos\theta + \beta_{p+2}\sin\theta + \epsilon$$

Multiple Imputation Procedure [Rubin, 1987] :

1. **Impute** - Use an imputation method g to impute Θ_{mis} by $\dot{\theta} = g(\mathbf{X}, \Theta_{obs})$ or $(\cos\dot{\theta}, \sin\dot{\theta})' = g(\mathbf{X}, \Theta_{obs})$ M times to create M completed data sets
2. **Analyze** - For each completed data $(\mathbf{X}, \Theta_{obs}, \Theta_{mis}^{(m)})$ for $m = 1, \ldots, M$, estimate β using least squares to collect $\hat{\beta}^{(m)}$ and $U^{(m)} = se(\hat{\beta}^{(m)})^2$
3. **Combine** - Apply **Rubin's rules** to get point estimate $\bar{Q}$ and total variance T

A Brief Intro to Directional Statistics

- Directional data consist of angles and points on the unit sphere

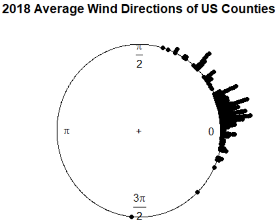


Figure 1: 2018 Average Wind Directions in US Counties

A Brief Intro to Directional Statistics

- Directional data consist of angles and points on the unit sphere
- Methods that use data unconstrained in $\mathbb{R}^p$ will be called *inline* and data that lie on S^1 will be called *angular/circular*

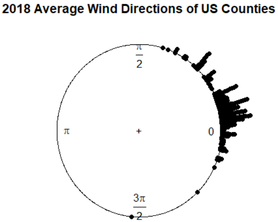


Figure 1: 2018 Average Wind Directions in US Counties

A Brief Intro to Directional Statistics

- Directional data consist of angles and points on the unit sphere
- Methods that use data unconstrained in $\mathbb{R}^p$ will be called *inline* and data that lie on S^1 will be called *angular/circular*
- $\theta \in [0, 2\pi)$ but $0 \equiv 2\pi$ and methods should be rotation invariant

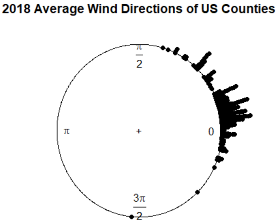


Figure 1: 2018 Average Wind Directions in US Counties

A Brief Intro to Directional Statistics

- Directional data consist of angles and points on the unit sphere
- Methods that use data unconstrained in $\mathbb{R}^p$ will be called *inline* and data that lie on S^1 will be called *angular/circular*
- $\theta \in [0, 2\pi)$ but $0 \equiv 2\pi$ and methods should be rotation invariant
- Distributions on the circle can be intrinsic (von Mises) or arise out of transformations (wrapping or projection)

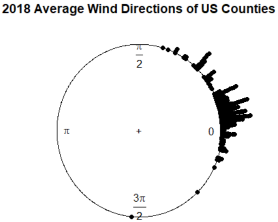


Figure 1: 2018 Average Wind Directions in US Counties

Incomplete Data Structure

Analysis model:

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \beta_{p+1} \cos\theta + \beta_{p+2} \sin\theta + \epsilon$$

$D = (\mathbf{X}, \Theta)$

$\mathbf{X}$	θ	$\cos\theta$	$\sin\theta$
$\mathbf{x}_1$	θ_1	$\cos\theta_1$	$\sin\theta_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{x}_{n_{obs}}$	$\theta_{n_{obs}}$	$\cos\theta_{n_{obs}}$	$\sin\theta_{n_{obs}}$
$\mathbf{x}_{n_{obs}+1}$	?	?	?
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{x}_n$	?	?	?

Θ_{obs}

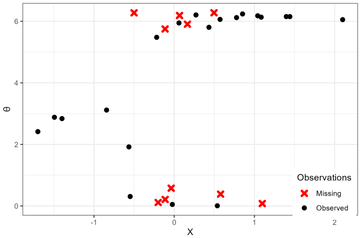
θ	$\cos\theta$	$\sin\theta$
θ_1	$\cos\theta_1$	$\sin\theta_1$
$\vdots$	$\vdots$	$\vdots$
$\theta_{n_{obs}}$	$\cos\theta_{n_{obs}}$	$\sin\theta_{n_{obs}}$

Θ_{mis}

θ	$\cos\theta$	$\sin\theta$
?	?	?
$\vdots$	$\vdots$	$\vdots$
?	?	?

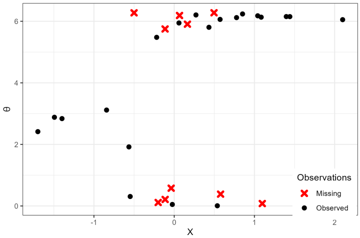
$\mathbf{X} = (X_1, \ldots, X_p, Y)$

Passive Imputation

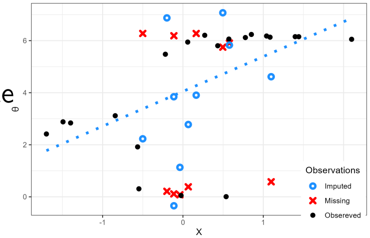


1. Impute by $\theta = \beta_0 + \beta_1 X + \beta_2 Y$
2. Transform by $\theta \mapsto (\cos\theta, \sin\theta)'$

Passive Imputation

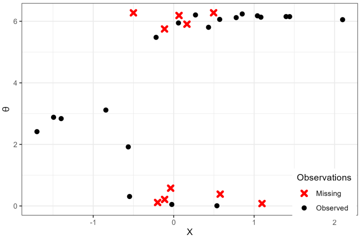


Impute

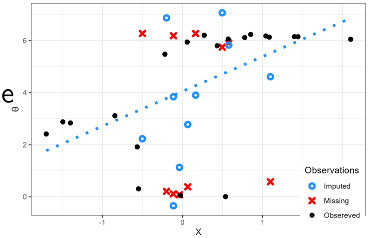


1. Impute by $\theta = \beta_0 + \beta_1 X + \beta_2 Y$
2. Transform by $\theta \mapsto (\cos\theta, \sin\theta)'$

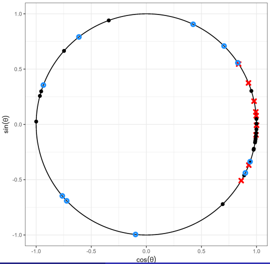
Passive Imputation



Impute



Transform

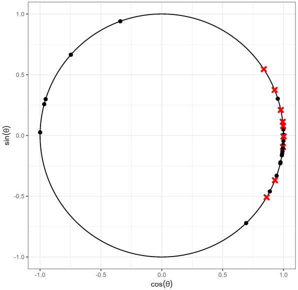


1. Impute by $\theta = \beta_0 + \beta_1 X + \beta_2 Y$
2. Transform by $\theta \mapsto (\cos\theta, \sin\theta)'$

Direct Imputation

- Impute cartesian coordinates instead by

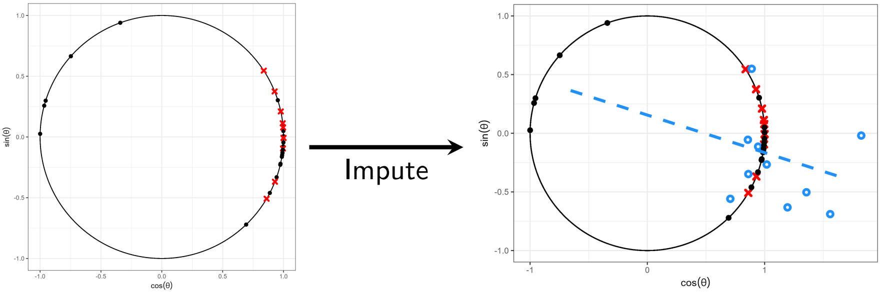
$$\begin{pmatrix} \cos\theta_i \\ \sin\theta_i \end{pmatrix} = \begin{pmatrix} \beta_{10} + \beta_{11} X_i + \beta_{12} Y_i \\ \beta_{20} + \beta_{21} X_i + \beta_{22} Y_i \end{pmatrix}$$



Direct Imputation

- Impute cartesian coordinates instead by

$$\begin{pmatrix} \cos\theta_i \\ \sin\theta_i \end{pmatrix} = \begin{pmatrix} \beta_{10} + \beta_{11}X_i + \beta_{12}Y_i \\ \beta_{20} + \beta_{21}X_i + \beta_{22}Y_i \end{pmatrix}$$



Congeniality

Congeniality The imputation should, at minimum, contain all of the variables that will be in the analysis model [Meng, 1994].

- With a circular model, imputing θ is equivalent to imputing $\mathbf{u}$ and any structure between θ and $Y, \mathbf{X}$ is in both the imputation and analysis models
- Imputing with an inline model, the imputation model is congenial if imputing on $\mathbf{u}$, but not if imputing on θ and then transforming to $\mathbf{u}$
  - Analysis model: $p(Y, X_1, \ldots, X_p, \cos\theta, \sin\theta)$
  - Uncongenial imputation model: $p_{Imp}(Y, X_1, \ldots, X_p, \theta)$ then transform $\theta \mapsto (\cos\theta, \sin\theta)'$
  - Congenial imputation model: $p_{Imp}(Y, X_1, \ldots, X_p, \cos\theta, \sin\theta)$

The Projected Normal Distribution I

Projected Normal Distribution

A multivariate normal vector $\mathbf{y}$ with $E(\mathbf{y}) = \mu$ and $Var(\mathbf{y}) = \Sigma$ normalized by its length results in a projected normal vector $\mathbf{u} = \mathbf{y}/||\mathbf{y}||$

- More flexible than the von Mises or wrapped family
- Equivalent to von Mises and wrapped normal when unimodal and symmetric for small and large concentrations
- Estimate latent lengths $l_i = ||\mathbf{y}_i||$ for each $\mathbf{u}_i \equiv \theta_i$ to work with $\tilde{\mathbf{y}}_i = \tilde{l}_i \mathbf{u}_i \dot{\sim} N_k(\mu, \Sigma)$ [Presnell et al., 1998, Nuñez-Antonio and Gutiérrez-Peña, 2005]

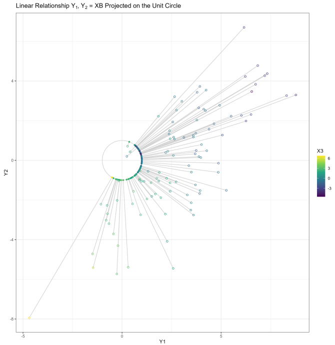
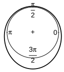


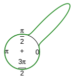
Figure 2: Projected normal regression with latent $\mathbf{y}$.

Projected Normal Examples

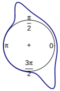
- If Σ is unconstrained, then the parameters are non-identifiable
  - Assume $\Sigma = I_k$ for symmetric and unimodal distribution
  - Assume some other constrained form to get bi-modal, skewed, or symmetric distributions [Hernandez-Stumpfhauser et al., 2017]



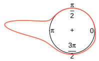
(a) Low concentration



(b) High concentration



(c) Bi-modal



(d) Skewed

Figure 3: Examples of projected normal distributions

Table of Contents

Standard Methods

- Let ω be one of $\{\theta, \cos\theta, \sin\theta\}$ depending on the imputation procedure with $\mathbf{X} = (1, Y, X_1, \ldots, X_p)$
- Assume only ω is incomplete

Standard Methods

- Let ω be one of $\{\theta, \cos\theta, \sin\theta\}$ depending on the imputation procedure with $\mathbf{X} = (1, Y, X_1, \ldots, X_p)$
- Assume only ω is incomplete

Linear Regression

- Use the observed data to fit the model $\omega_i = \mathbf{x}_i'\gamma + \eta_i$ for $i = 1, \ldots, n_{obs}$; $\eta_i \overset{iid}{\sim} N(0, \tau^2)$
- Impute ω_i by drawing from the predictive posterior of the regression model given the predictors $\mathbf{x}_j$ for each incomplete observation, $j \notin \{1, \ldots, n_{obs}\}$

Standard Methods

- Let ω be one of $\{\theta, \cos\theta, \sin\theta\}$ depending on the imputation procedure with $\mathbf{X} = (1, Y, X_1, \ldots, X_p)$
- Assume only ω is incomplete

Linear Regression

- Use the observed data to fit the model $\omega_i = \mathbf{x}_i'\gamma + \eta_i$ for $i = 1, \ldots, n_{obs}$; $\eta_i \overset{iid}{\sim} N(0, \tau^2)$
- Impute ω_i by drawing from the predictive posterior of the regression model given the predictors $\mathbf{x}_j$ for each incomplete observation, $j \notin \{1, \ldots, n_{obs}\}$

Predictive Mean Matching

- Fit a Bayesian regression with a weakly informative prior
- Use draws from the predictive posterior to create a set of neighbors for drawing imputations

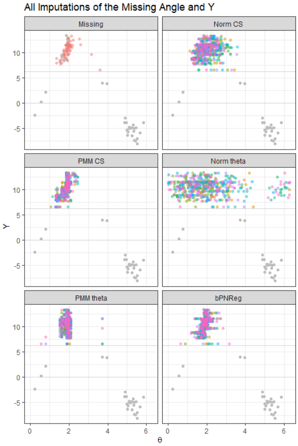
Projected Normal Regression I

- Let $\theta_i \sim PN_2(\mu_i, I_2)$ where $\mu_i = \mathbf{x}_i'\mathbf{B}$ and $\mathbf{B} = (\beta_1, \beta_2) \in \mathbb{R}^{(p+2)\times 2}$
- Normal prior on $\beta_{ij} \sim N(0, 10000)$
    - Implemented by `bpnreg` package in R [Cremers, 2021]
- Assumes a symmetric and unimodal distribution for θ_i given $\mathbf{X}_i$

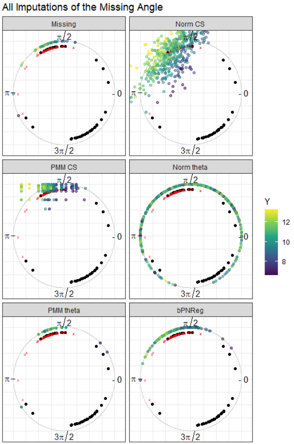
Projected Normal Imputation

- Fit the projected normal regression model using the observed data and obtain posterior draws $\dot{\mathbf{B}}$ of $\mathbf{B}$
- Sample $\dot{Y}_j \sim N_2(\mathbf{x}_j'\dot{\mathbf{B}}, I_2)$ and normalize to get $U_j = Y_j/||Y_j||^{-1} = (\cos\dot{\theta}_j, \sin\dot{\theta}_j)'$ for $j \notin \{1, \ldots, n_{obs}\}$
- Impute the missing θ_j with predictive posterior draws of $\dot{\theta}_j$

Projected Normal Regression II



(a) On the unit circle



(b) Coordinates in $\mathbb{R}^2$

Figure 4: Imputations with angular data generated by projected normal regression

Table of Contents

Simulation Overview

How well do the projected normal imputations perform under various settings?

1. Different Sample Sizes - $N = 50; 100; 500; 1,000$

Simulation Overview

How well do the projected normal imputations perform under various settings?

1. Different Sample Sizes - $N = 50; 100; 500; 1,000$
2. Different Missing Data Proportions - With $N = 100$, $p_{miss} = 0.1, 0.5, 0.9$

Simulation Overview

How well do the projected normal imputations perform under various settings?

1. Different Sample Sizes - $N = 50; 100; 500; 1,000$
2. Different Missing Data Proportions - With $N = 100$, $p_{miss} = 0.1, 0.5, 0.9$
3. Different Data Generating Processes
   a. High Concentration Projected Normal
   b. Low Concentration Projected Normal
   c. Skewed Projected Normal
   d. Bi-modal Projected Normal
   e. Projected Normal Regression
   f. von Mises Regression
   $\theta \sim M(\mu_i, 15)$ with $\mu_i = 2\arctan(-3X_{i1} + 3X_{i2})$

Results I

- Projected normal imputation provides better inferences (nominal CI coverage with low average width; low bias and MSE) compared to imputations of just θ using linear regression or PMM
- Imputations on $(\cos\theta, \sin\theta)'$ also provide robust inferences despite not being angular
- All investigated imputation strategies perform similarly across differing sample sizes and missingness proportions; wider intervals/greater MSE with low N and high p_{miss}

Results II

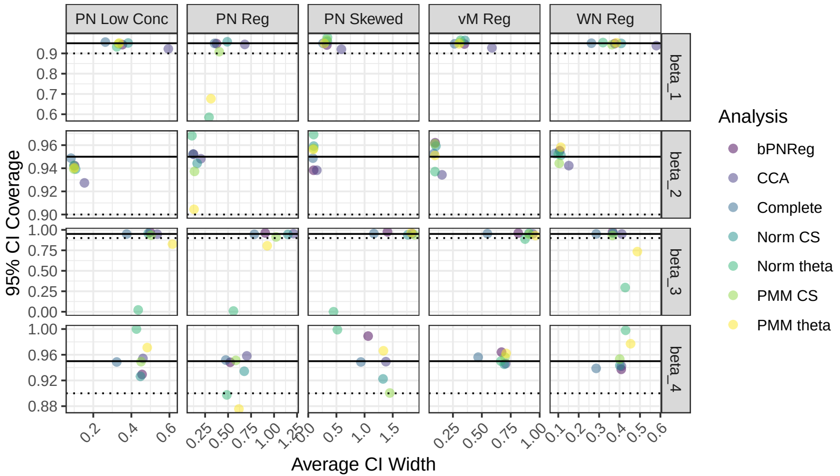


Figure 5: 95% Confidence interval coverages and average widths for each coefficient using the different analyses and data generating processes.

Table of Contents

Overview I

- Air pollution is strongly connected to asthma [Glad et al., 2012, Gleason et al., 2014, Lin et al., 2002, Weisel et al., 1995, Wilson et al., 2005]
- Model the impact of air pollution on asthma accounting for meteorological conditions
- Data collected at the US county-level and averaged over the year from CDC () and EPA (2021)
- Census data (2021) is used to weight the county observations by population

Overview II

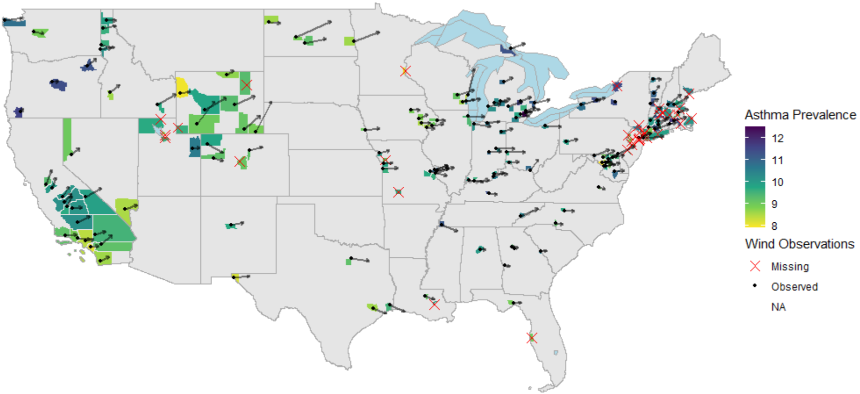


Figure 6: Map of US counties asthma prevalence and average wind directions and speed.

Analysis

- Impute using each of the previously discussed methods for the angular data
- Impute inline data with linear regression or predictive mean matching
- Include all variables collected for this step

Analysis

- Impute using each of the previously discussed methods for the angular data
- Impute inline data with linear regression or predictive mean matching
- Include all variables collected for this step

- Mixed effects linear model (centered and scaled predictors)
  - Fixed effects for the meteorological and air pollution variables
  - Random intercepts for the state-level

$$
\begin{aligned}
Asthma_i = {} & \beta_0 + \beta_1 AirPressure_i + \beta_2 RH_i + \beta_3 Temperature_i \\
& + \beta_4 WindSpeed_i + \beta_5 \cos WindAngle_i + \beta_6 \sin WindAngle_i \\
& + \beta_7 USGD_i + \beta_8 MedianAQI_i + \beta_9 90thPercentileAQI_i \\
& + \beta_{10} CODays_i + \beta_{11} NO2Days_i + \beta_{12} Ozone_i \\
& + \beta_{13} PM2.5_i + \beta_{14} PM10_i + \mathbf{Z}\eta + \epsilon_i
\end{aligned}
$$

Results I

- Imputation strategies are generally consistent and provide better estimates than complete case analysis (CCA)
- Projected normal regression and linear regression imputations on $(\cos\theta, \sin\theta)$ produce similar inferences
- Predictive mean matching with either θ or $(\cos\theta, \sin\theta)'$ are similar

Results II

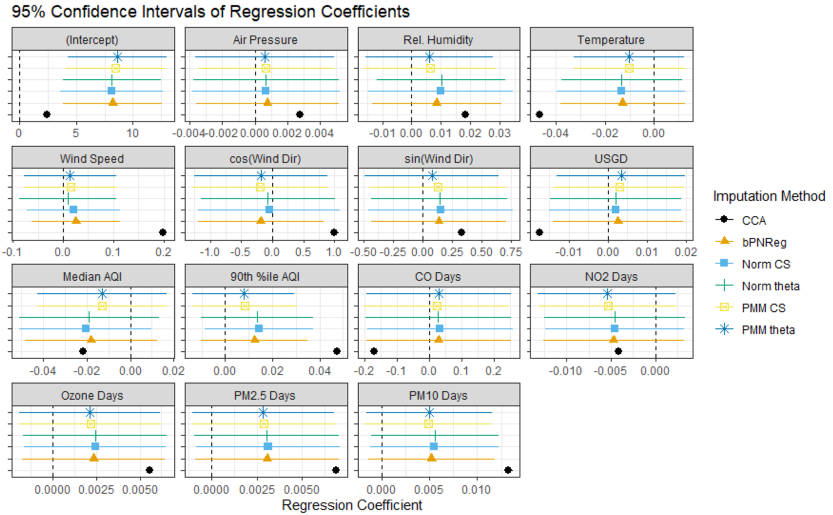


Figure 7: Bias of coefficient estimates using different imputation strategies.

Table of Contents

What's Next?

- Imputation for spherical data
- Applying angular imputation models when the response or outcome is angular
- Incomplete data analysis for angular data in temporal or spatial settings

Acknowledgements

- Research was supported by NSF AGEP-GRS supplement for Award #2015320.
- Thank you to my advisor Dr. Ofer Harel for support and assistance throughout the research process.

Questions?

References I

US Census Bureau. County Population Totals: 2010-2019. https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html, 2021.

Centers for Disease Control CDC and Prevention. National Environmental Public Health Tracking Network Data Explorer. https://ephtracking.cdc.gov/DataExplorer/?c=11. Accessed: 10/4/2022.

Jolien Cremers. Bpnreg: Bayesian Projected Normal Regression Models for Circular Data, August 2021.

Jo Ann Glad, LuAnn Lynn Brink, Evelyn O. Talbott, Pei Chen Lee, Xiaohui Xu, Melissa Saul, and Judith Rager. The Relationship of Ambient Ozone and PM2.5 Levels and Asthma Emergency Department Visits: Possible Influence of Gender and Ethnicity. *Archives of Environmental & Occupational Health*, 67(2):103–108, April 2012. ISSN 1933-8244. doi: 10.1080/19338244.2011.598888.

References II

Jessie A. Gleason, Leonard Bielory, and Jerald A. Fagliano. Associations between ozone, PM2.5, and four pollen types on emergency department pediatric asthma events during the warm season in New Jersey: A case-crossover study. *Environmental Research*, 132:421–429, July 2014. ISSN 1096-0953. doi: 10.1016/j.envres.2014.03.035.

Ofer Harel and Xiao-Hua Zhou. Multiple imputation: Review of theory, implementation and software. *Statistics in Medicine*, 26(16):3057–3077, July 2007. ISSN 02776715, 10970258. doi: 10.1002/sim.2787.

Daniel Hernandez-Stumpfhauser, F. Jay Breidt, and Mark J. van der Woerd. The General Projected Normal Distribution of Arbitrary Dimension: Modeling and Bayesian Inference. *Bayesian Analysis*, 12(1): 113–133, March 2017. ISSN 1936-0975, 1931-6690. doi: 10.1214/15-BA989.

References III

Mei Lin, Yue Chen, Richard T. Burnett, Paul J. Villeneuve, and Daniel Krewski. The influence of ambient coarse particulate matter on asthma hospitalization in children: Case-crossover and time-series analyses. *Environmental Health Perspectives*, 110(6):575–581, June 2002. ISSN 0091-6765. doi: 10.1289/ehp.02110575.

Xiao-Li Meng. Multiple-Imputation Inferences with Uncongenial Sources of Input. *Statistical Science*, 9(4):538–558, 1994. ISSN 0883-4237.

Gabriel Nuñez-Antonio and Eduardo Gutiérrez-Peña. A Bayesian analysis of directional data using the projected normal distribution. *Journal of Applied Statistics*, 32(10):995–1001, December 2005. ISSN 0266-4763. doi: 10.1080/02664760500164886.

Brett Presnell, Scott P. Morrison, and Ramon C. Littell. Projected Multivariate Linear Models for Directional Data. *Journal of the American Statistical Association*, 93(443):1068–1077, September 1998. ISSN 0162-1459. doi: 10.1080/01621459.1998.10473768.

References IV

Donald B Rubin. *Multiple Imputation for Nonresponse in Surveys | Wiley Series in Probability and Statistics*. Wiley series in probability and mathematical statistics : Applied probability and statistics. Wiley, New York, June 1987. ISBN 978-0-471-08705-2.

OAR US Environmental Protection Agency. Air Quality System Data Mart [internet database]. https://www.epa.gov/outdoor-air-quality-data, August 2021.

C. P. Weisel, R. P. Cody, and P. J. Lioy. Relationship between summertime ambient ozone levels and emergency department visits for asthma in central New Jersey. *Environmental Health Perspectives*, 103 Suppl 2 (Suppl 2):97–102, March 1995. ISSN 0091-6765. doi: 10.1289/ehp.95103s297.

References V

Adam M. Wilson, Cameron P. Wake, Tom Kelly, and Jeffrey C. Salloway. Air pollution, weather, and respiratory emergency room visits in two northern New England cities: An ecological time-series study. *Environmental Research*, 97(3):312–321, March 2005. ISSN 0013-9351. doi: 10.1016/j.envres.2004.07.010.

Appendix I

Rubin's Rules - Multivariate [Rubin, 1987, Harel and Zhou, 2007]

- $\bar{Q} = \frac{1}{M} \sum_{m=1}^{M} \hat{\beta}^{(m)}$ is the mean of the estimates
- $\bar{U} = \frac{1}{M} \sum_{m=1}^{M} U^{(m)}$ is the mean of the within variance
- $B = \frac{1}{M-1} \sum_{m=1}^{M} (\hat{\beta}^{(m)} - \bar{Q})(\hat{\beta}^{(m)} - \bar{Q})'$ is the between variance
- $T = \bar{U} + (1 + \frac{1}{M})B$ is the total variance
- Proportionality assumption: $B \propto \bar{U}$ so that $\tilde{T} = (1-t)\bar{U}$ where $t = (1 + \frac{1}{M}) tr(B\bar{U}^{-1})$

Appendix II

Simulated Data Generating Processes

Different Data Generating Processes parametrized by Φ_j

1. High Concentration Projected Normal - $\Phi_1 = ((10, 0)', I_2)'$
2. Low Concentration Projected Normal - $\Phi_2 = ((1, 0)', I_2)'$
3. Skewed Projected Normal - $\Phi_3 = \left(\begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} \right)'$
4. Bi-modal Projected Normal - $\Phi_4 = \left(\begin{pmatrix} 0.5 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ -2 & 10 \end{pmatrix} \right)'$
5. Projected Normal Regression - $\Phi_5 = (\mathbf{B}, I_2)'$ so that $\mu_i' = (1, X_{i1}, X_{i2})\mathbf{B}$ and $\mathbf{B} = (\beta_1, \beta_2)'$ where $\beta_1 = (0, 1, -0.5)'$ and $\beta_2 = (0, -3, 1)'$
6. von Mises Regression - $\Phi_6 = (0, -3, 3, 15)'$ so that $\theta \sim M(\mu_i, 15)$ with $\mu_i = 2\arctan(-3X_{i1} + 3X_{i2})$

Appendix III

Simulation Procedure

1. Create data set and impose missing values via logistic model
   a. $(X_1, X_2)' \sim N_2\left(\begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ -2 & 10 \end{pmatrix}\right)$
   b. $\theta \sim F(\theta | X_1, X_2, \Phi)$ where F is the projected normal or von Mises distributions/regressions parametrized by Φ
   c. $Y = \mathbf{X}\beta + \epsilon$ where $\epsilon \sim N(0, 1)$
2. Impute the missing angular data using each method M times
3. Fit the regression model $Y = \mathbf{X}\beta$ to each imputed data set, the complete data, and complete case analysis
4. Combine results for multiply imputed data sets using Rubin's rules
5. Repeat 1000 times

Appendix IV

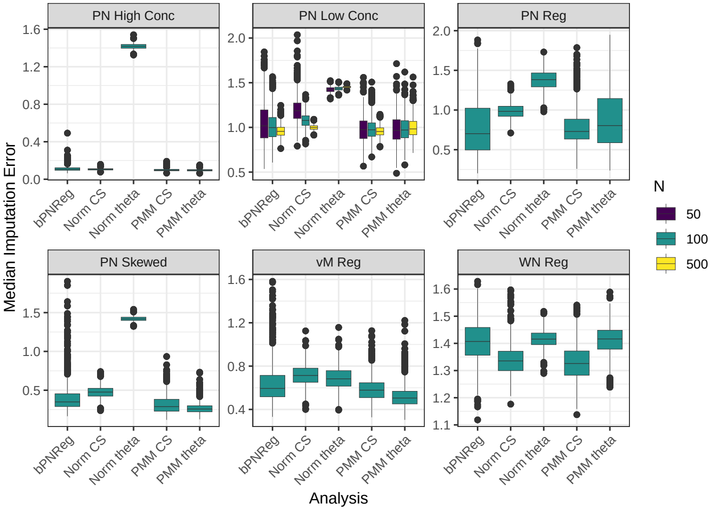


Figure 8: Median imputation error for p(miss) = 0.5.

Appendix V

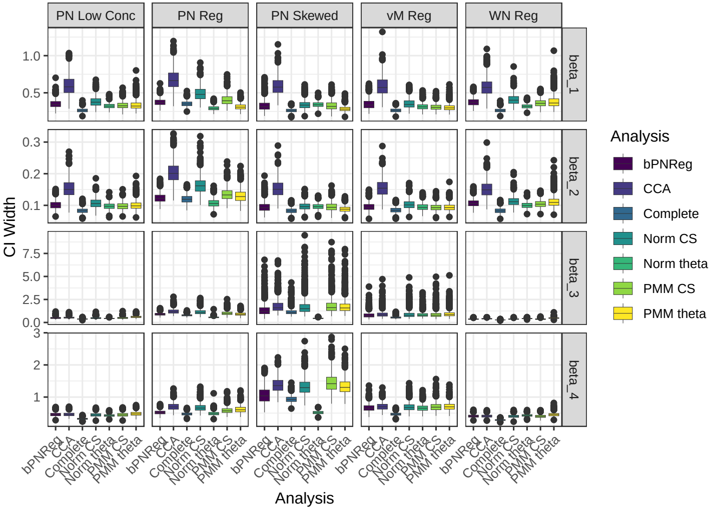
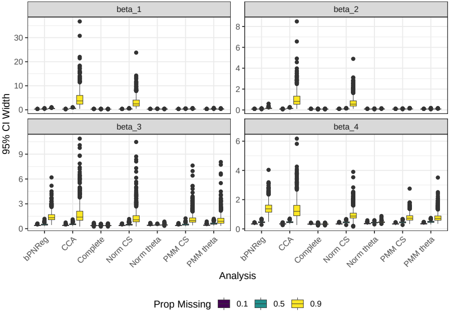


Figure 9: 95% Confidence interval widths for each coefficient using the different analyses and data generating processes.

Appendix VI



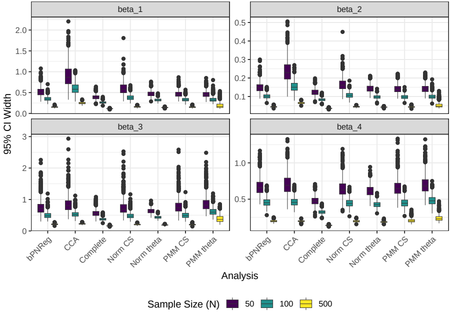


Figure 10: Varying p_{miss} with $N = 100$. Figure 11: Varying N with $p_{miss} = 0.5$.

95% Confidence interval widths for each coefficient with low concenctration projected normal angular data.